JAMA Network Open

Original Investigation | Public Health

Gender-Affirming Hormone Therapy and Depressive Symptoms Among Transgender Adults

Sari L. Reisner, ScD, MA; David R. Pletta, PhD, MPH; Alex S. Keuroghlian, MD, MPH; Kenneth H. Mayer, MD; Madeline B. Deutsch, MD, MPH; Jennifer Potter, MD; Jaclyn M. W. Hughto, PhD, MPH; Alexander Harris, MPH; Asa E. Radix, MD, PhD, MPH

Abstract

IMPORTANCE In the US, transgender, nonbinary, and gender diverse (TGD) adults have high rates of depression. Gender-affirming hormone therapy (GAHT) is associated with improved mental health outcomes, yet existing US studies have short follow ups and lack sample diversity.

OBJECTIVE To evaluate the use of GAHT delivered in primary care as an intervention for moderate-to-severe depressive symptoms in diverse TGD adult patients.

DESIGN, SETTING, AND PARTICIPANTS LEGACY was an observational cohort study conducted in federally qualified community health centers in Boston and New York that followed up TGD patients (N = 3592) from calendar years 2016 to 2019 (48 months). Participants included individuals aged 18 years or older, gender identity different from sex at birth, a past 12-month medical visit, and signed patient consent form in the electronic health record (EHR).

EXPOSURES Prescriptions for GAHT obtained from EHR data, using the date of the first and last GAHT prescription in each calendar year of observation (GAHT within the year vs no GAHT during the year).

MAIN OUTCOMES AND MEASURES A binary outcome of patient-reported moderate-to-severe depressive symptoms was obtained using the validated Patient-Health Questionnaire (PHQ), scoring 10 or greater on the PHQ-9 or scoring 3 or greater on the PHQ-2. Following multiple imputation, generalized estimating equations (GEE) longitudinally modeled GAHT and moderate-to-severe depressive symptoms (n = 20 320 observations) and adjusted for age, gender identity, race and ethnicity, health insurance, federal poverty level, HIV serostatus, number of cohort years, and clinical site.

RESULTS The median age of the 3592 patients was 28 (IQR, 24-36) years. Race and ethnicity was diverse (1.3% Asian/Pacific Islander, 11.7% Black, 16.1% Hispanic/Latinx, 63.1% White, 6.8% multiracial, and 1.4% other). In addition, 18.9% were nonbinary, 52.1% lived below the federal poverty level, 34.2% were publicly insured, 4.1% were uninsured, and 5.1% were living with HIV. At baseline, 84.5% of the individuals were prescribed GAHT and 15.3% reported moderate-to-severe depressive symptoms. Patients prescribed GAHT had a statistically significantly lower risk of moderate-to-severe depressive symptoms over follow-up compared with those not prescribed GAHT (adjusted risk ratio, 0.85; 95% CI, 0.75-0.98).

CONCLUSIONS AND RELEVANCE In this longitudinal observational cohort study, GAHT was associated with lower rates of moderate-to-severe depressive symptoms, highlighting the importance of gender-affirming primary care models for TGD patients.

Key Points

Question Is gender-affirming hormone therapy (GAHT) associated with lower rates of moderate to severe depressive symptoms in transgender, nonbinary, and gender diverse (TGD) adults in primary care?

Findings In this cohort study of 3592 TGD patients in primary care at federally qualified community health centers in Boston and New York, 15.3% had moderate-to-severe depressive symptoms assessed using the Patient Health Questionnaire (PHQ). In addition, GAHT was associated with a lower risk of moderate-to-severe depressive symptoms across 48 months of follow-up.

Meaning The findings of this study suggest that integrated GAHT with primary care and low-barrier GAHT access is associated with lower rates of mental health morbidity in TGD patients.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

JAMA Network Open. 2025;8(3):e250955. doi:10.1001/jamanetworkopen.2025.0955

Open Access. This is an open access article distributed under the terms of the CC-BY-NC-ND License, which does not permit alteration or commercial use, including those for text and data mining, AI training, and similar technologies.

Downloaded from jamanetwork.com by guest on 03/18/2025

Introduction

In the US, major depression is a leading cause of morbidity and mortality,[1-3] affecting an estimated 21.0 million adults aged 18 years or older (8.3% of all US adults).[3] In 2019, US national probability sample surveillance data found that 6.8% of adults met the criteria for moderate-to-severe depressive symptoms, using the 8-item Patient-Health Questionnaire (PHQ-8).[4] The US Preventive Services Task Force recommends routine screening for depression in the general adult population aged 18 years or older in primary care settings, using tools such as the PHQ, to identify those who should be further diagnostically evaluated for depression and referred to mental health services.[5-7]

Transgender, nonbinary, and gender diverse (TGD) adults in the US experience health disparities, including a 2- to 4-fold increased risk of depression relative to cisgender adults.[8,9] Gender-affirming hormone therapy (GAHT) is known to be medically necessary and improve mental health outcomes.[10-13] Yet, many studies are cross-sectional, and longitudinal studies assessing the GAHT and mental health outcomes often rely on samples from specialized gender clinics, are homogeneous in sociodemographic characteristics, have relatively short follow-up periods, and do not consider the context of other TGD health disparities,[10-13] such as high HIV burden.[14] Data are lacking in clinical samples of TGD primary care settings and with diversity in terms of age, gender identity, race and ethnicity, socioeconomic characteristics, and HIV status. Furthermore, data are needed from clinical settings that implement GAHT using the informed consent protocol, which provides greater autonomy to patients in GAHT decision-making compared with historical GAHT protocols that required involvement of mental health professionals,[15] to understand the effectiveness of this model of care delivery for TGD adults.

This study sought to evaluate whether GAHT was associated with 12-month improvements in past 2-week moderate-to-severe depressive symptoms for diverse TGD primary care patients over 48 months of follow-up. The hypothesis was that patients prescribed GAHT would have lower rates of moderate-to-severe depressive symptoms compared with those not prescribed GAHT.

Methods

Participants and Procedures

LEGACY was a longitudinal multisite cohort study of TGD adult primary care patients from 2 federally qualified community health centers (FQHCs) in Boston, Massachusetts (Fenway Health) and New York, New York (Callen-Lorde). Both clinical sites had long histories providing GAHT for TGD patients using an informed consent model, reducing unnecessary barriers or gatekeeping practices related to mental health diagnosis or treatment requirements for GAHT.[15,16] Eligibility criteria were age 18 years or older, gender identity different from the sex assigned at birth (assessed gender identity and sex using the 2-step method from patient registration forms, and/or *International Statistical Classification of Diseases and Related Health Problems, 10th Revision* codes F64.0-F64.9), a primary care patient defined as having a past 12-month medical visit, and signed patient informed consent form acknowledging that patient data may be used for research with no research exclusion in the EHR. LEGACY patients were not individually compensated; deidentified EHR data were extracted. Formative qualitative research informed the study[17,18] and a community advisory board provided input. The detailed cohort method has been published.[19] This analysis was restricted to LEGACY patients who completed a baseline screener for depressive symptoms in 2016 (N = 3592) and at least 1 additional depressive symptom screener across follow-up between calendar years 2016 and 2019. Use of GAHT in primary care was delivered according to the same standards by Fenway Health and Callen-Lorde.[15] Study procedures were approved by the Fenway Health Institutional Review Board. This report follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.[20]

Downloaded from jamanetwork.com by guest on 03/18/2025

Measures

Clinical Outcome: Moderate-to-Severe Depressive Symptoms

Patient-reported moderate-to-severe depressive symptoms in the past 2 weeks vs no such symptoms were assessed using the validated PHQ,[21,22] completed in routine care at both sites. Patients were asked questions such as "Over the last 2 weeks, how often have you been bothered by the following problems?" (response options: not at all = 0, several days = 1, more than half the days = 2, nearly every day = 3). Fenway Health used the 9-item version (PHQ-9) with high reliability (α = 0.89) and scores ranging from 0 to 27. The validated cut-point for moderate to severe depression at a PHQ-9 score of 10 or above (88% sensitivity, 88% specificity for major depression) was selected as suggesting clinically significant depressive distress.[21] Callen-Lorde administered the 2-item version (PHQ-2), with scores ranging from 0 to 6, and a PHQ-2 score of 3 or above suggestive of possible major depression (83% sensitivity, 92% specificity).[22] Screening positive for clinically significant depressive symptoms (hereinafter moderate-to-severe depressive symptoms) used the validated cut-points for PHQ-9 or PHQ-2.

A binary operationalization of moderate-to-severe depressive symptoms vs no such symptoms was used to ensure consistency with the US Preventive Services Task Force recommendation of routine screening for depression in the general adult population in primary care settings, using tools such as the PHQ with clinically relevant cut-points, to identify those who warrant additional diagnostic evaluation for depression and referrals to mental health services.[5-7]

Exposure: GAHT Prescription

The exposure was a binary variable of being prescribed GAHT in EHR data, operationalized using the date of the first GAHT prescription and the date of the last GAHT prescription in each calendar year of observation (GAHT within the year vs no GAHT during the year). Hormone prescription types were pubertal blockers (eg, leuprolide), antiandrogen (eg, spironolactone), estrogen (eg, estradiol), progesterone (eg, micronized progesterone), and testosterone (eg, testosterone gel or injection).

Covariates and Confounders

Baseline age was derived from age within the year of the observation (ages 18-24, 25-30, 31-40, 41-50, or ≥51 years). Assigned sex at birth and current gender identity were assessed at patient registration and coded woman/transgender women, man/transgender man, nonbinary assigned male sex at birth, nonbinary assigned female sex at birth, and another transgender identity not specified. Self-reported race in the EHR was American Indian or Alaska Native, Asian, Black or African American (hereinafter, Black), Multiracial, Native Hawaiian or Other Pacific Islander, White, and another racial identity. Hispanic or Latinx ethnicity was assessed separately as a binary variable. Data on race and ethnicity were included because it is associated with depression in this population. Health insurance was private, public, or uninsured. The federal poverty level used US benchmarks of 0% to 138%, 139% to 250%, and greater than or equal to 251% of the poverty level based on the sample distribution and aligned with the baseline year.[23] Cohort years (2016-2019) ranged from 1 to 4. Site was coded as a binary variable for Fenway Health or Callen-Lorde.

The acquisition and transmission risk of HIV was based on laboratory-confirmed HIV serostatus and prevention-effective adherence. Prevention-effective adherence refers to aligning consistent preexposure prophylaxis (PrEP) use or viral suppression coverage with periods of HIV risk.[24,25] For HIV-negative patients, adherence refers to past 12-month PrEP uptake (a PrEP prescription for emtricitabine-tenofovir disoproxil fumarate) within the last 12 months vs none. For patients living with HIV, adherence refers to past 12-month virological suppression, derived from laboratory-confirmed viral load suppression (<200 copies/mL) vs not (≥200 copies/mL) based on the most recent viral load test in the last 12 months and consistent with clinical guidelines.[25] Laboratory-confirmed HIV serostatus, PrEP prescription, and viral suppression variables were combined into a 4-category composite: HIV-negative, not prescribed PrEP; HIV-negative, prescribed PrEP; living with HIV, not virally suppressed; and living with HIV, virally suppressed.

Downloaded from jamanetwork.com by guest on 03/18/2025

Statistical Analysis

Statistical analyses were implemented in RStudio, version 2022.07.1 and R, version 4.2.0 (R Foundation for Statistical Computing). Descriptive statistics (frequencies, median, IQR) were calculated for the baseline cohort overall (N = 3592) and stratified by moderate-to-severe depressive symptoms. Missing data assumptions were assessed via a modularization approach for EHR-based studies.[26] Missingness ranged from 0.2% (age in years) to 14.1% (Latinx ethnicity). Linear models were fit to assess estimators of attrition. Patients with moderate-to-severe depressive symptoms at baseline in 2016 had an average of 0.17 more observation between 2016 and 2019 (P = .002) compared with those with no moderate-to-severe depressive symptoms at baseline, meaning they were more likely to be retained in care (greater number of years in care). The monotonic missing data pattern resulting from loss to follow-up was addressed using multiple imputation procedures. Multiple imputation was used to construct 5 analytic datasets for statistical modeling.[27] Imputed datasets were generated using the MICE package (multivariate imputation via chained equations) in R.[28]

For longitudinal analyses, log-Poisson generalized estimating equations (GEE) with cluster-robust SEs[29,30] and an exchangeable working correlation structure (selected using the quasi-likelihood under independence model criterion statistic) were used to estimate the marginal or population-averaged outcome of concurrent GAHT (exposure) and moderate-to-severe depressive symptoms, adjusting for age, gender identity, race and ethnicity, poverty level, health insurance, HIV acquisition and transmission risk, clinical site, and cohort years. Covariate-adjusted GEEs were fit across each of the imputed datasets, and results were pooled using Rubin rules.[31] Risk ratios (RRs) and 95% CIs were estimated. Generalized estimating equation models were selected due to their efficiency in estimation of regression parameters, resilience to model misspecification, and conservative approach to estimation of longitudinal effects.[32] A sensitivity analysis was conducted to assess how robust the treatment-effect estimate was to potential unmeasured or uncontrolled confounding. We calculated the E-value necessary, on the relative risk scale, to explain the observed association between GAHT prescription and the risk of moderate-to-severe depressive symptoms.[33,34] A larger E-value indicates a treatment effect estimate is more robust to unmeasured confounding. Statistical significance was set at the α = .05 level, with 2-sided significance tests.

Results

Descriptive Sample Characteristics

Table 1 presents baseline characteristics of the 3592 patients overall and stratified by moderate-to-severe depressive symptoms: Median age was 28 (IQR, 24-36) years. the study population was diverse (4.4% Asian, 11.7% Black, 16.1% Hispanic/Latinx, 0.8% Native Hawaiian or Other Pacific Islander, 63.1% White, 6.8% Multiracial, and 1.4% another racial identity); 18.9% were nonbinary; 52.1% were socioeconomically marginalized (living 0%-138% federal poverty level); 34.2% had public insurance, and 4.1% had no insurance. Overall, 5.1% of the patients were living with HIV, with 22.0% not virally suppressed; 1.7% of HIV-negative patients were prescribed PrEP.

The prevalence of GAHT prescription and moderate-to-severe depressive symptoms from 2016 to 2019 is reported in **Table 2**. At baseline, 84.5% of the patients were prescribed GAHT and 15.3% had moderate-to-severe depressive symptoms. Across follow-up, the proportion of patients with a GAHT prescription ranged from 92.3% to 94.2%, and moderate-to-severe depressive symptoms ranged from 11.5% to 14.6%.

Longitudinal Modeling of GAHT and Moderate-to-Severe Depressive Symptoms

The number of observations was longitudinally modeled (N = 25 320) and are displayed in **Table 3**. Patients prescribed GAHT had a significantly lower risk of moderate-to-severe depressive symptoms over the follow-up period compared with those not prescribed GAHT (adjusted RR [ARR], 0.85; 95% CI, 0.75-0.98).

Downloaded from jamanetwork.com by guest on 03/18/2025

Table 1. Cohort Characteristics of Transgender, Nonbinary, and Gender-Diverse Patients: Overall and Stratified by PHQ Depression at Baseline in 2016

Characteristic	Patients, No. (%)		
	Total sample (N = 3592)	PHQ depression[a] (n = 551)	No PHQ depression (n = 3041)
Exposure			
Prescribed gender-affirming hormones			
Yes	3034 (84.5)	442 (80.2)	2592 (85.2)
No	558 (15.5)	109 (19.8)	449 (14.8)
Covariates and confounders			
Site			
Callen-Lorde	1290 (35.9)	80 (14.5)	1210 (39.8)
Fenway Health	2302 (64.1)	471 (85.5)	1831 (60.2)
Age group, y[b]			
18-24	1044 (29.1)	191 (34.7)	853 (28.0)
25-30	1068 (29.7)	168 (30.5)	900 (29.6)
31-40	802 (22.3)	105 (19.1)	697 (22.9)
41-50	335 (9.3)	48 (8.7)	287 (9.4)
≥51	337 (9.4)	39 (7.1)	298 (9.8)
Missing	6 (0.2)	0	6 (0.2)
Gender identity			
Transgender woman	1270 (35.4)	213 (38.7)	1057 (34.8)
Transgender man	1171 (32.6)	165 (29.9)	1006 (33.1)
Nonbinary AFAB	451 (12.6)	93 (16.9)	358 (11.8)
Nonbinary AMAB	227 (6.3)	39 (7.1)	188 (6.2)
Transgender, gender not specified	473 (13.2)	41 (7.4)	432 (14.2)
Racial identity[c]			
American Indian or Alaskan Native	17 (0.5)	3 (0.5)	14 (0.5)
Asian	158 (4.4)	19 (3.4)	139 (4.6)
Black or African American	419 (11.7)	44 (8.0)	375 (12.3)
Multiracial	244 (6.8)	40 (7.3)	204 (6.7)
Native Hawaiian or Other Pacific Islander	28 (0.8)	1 (0.2)	27 (0.9)
White	2268 (63.1)	385 (69.9)	1883 (61.9)
Another racial identity	49 (1.4)	10 (1.8)	39 (1.3)
Missing	409 (11.4)	49 (8.9)	360 (11.8)
Ethnicity			
Latinx or Hispanic	579 (16.1)	64 (11.6)	515 (16.9)
Not Latinx or Hispanic	2507 (69.8)	426 (77.3)	2081 (68.4)
Missing	506 (14.1)	61 (11.1)	445 (14.6)
Insurance			
Private	1983 (55.2)	291 (52.8)	1692 (55.6)
Public	1228 (34.2)	224 (40.7)	1004 (33.0)
Uninsured	149 (4.1)	27 (4.9)	122 (4.0)
Missing	232 (6.5)	9 (1.6)	223 (7.3)
Federal poverty level, %			
0-138	1871 (52.1)	326 (59.2)	1545 (50.8)
139-250	455 (12.7)	69 (12.5)	386 (12.7)
≥251	1021 (28.4)	130 (23.6)	891 (29.3)
Missing	245 (6.8)	26 (4.7)	219 (7.2)
HIV risk category			
HIV-negative, not prescribed PrEP	3336 (92.9)	511 (92.7)	2825 (92.9)
HIV-negative, prescribed PrEP	59 (1.6)	16 (2.9)	43 (1.4)
Living with HIV, virally suppressed[d]	142 (4.0)	18 (3.3)	124 (4.1)
Living with HIV, not virally suppressed	40 (1.1)	5 (0.9)	35 (1.2)
Missing	15 (0.4)	1 (0.2)	14 (0.5)
Total years in cohort, mean (SD)	3.20 (1.07)	3.15 (1.09)	3.21 (1.07)

Abbreviations: AFAB, assigned female sex at birth; AMAB, assigned male sex at birth; PHQ, Patient Health Questionnaire; PrEP, preexposure prophylaxis.

[a] Depression was defined as having a PHQ-2 score of 3 points or more (for Callen-Lorde patients) or a PHQ-9 score of 10 points or more (for Fenway Health patients) indicative of clinically significant symptoms of distress in the past 2 weeks (moderate-to-severe depressive symptoms).

[b] The median age of the cohort was 28 (IQR, 24-36) years.

[c] Race and ethnicity are associated with depression and were adjusted for in analyses.

[d] Virally suppressed is defined as having an HIV viral load count <200 copies/mL. Overall, 78.0% of patients living with HIV were virally suppressed; 1.7% of HIV-negative patients were prescribed PrEP.

Downloaded from jamanetwork.com by guest on 03/18/2025

Several covariates and confounders also evidenced associations with moderate-to-severe depressive symptoms (Table 3). Risk for moderate-to-severe depressive symptoms was higher for transgender women (ARR, 1.18; 95% CI, 1.05-1.32) and nonbinary assigned female sex at birth (ARR, 1.17; 95% CI, 1.01-1.35) vs transgender men; patients with public vs private insurance (ARR, 1.35; 95% CI, 1.22-1.49); for the lowest federal poverty level groups vs highest group (0-138% vs ≥251%: ARR, 1.26; 95% CI, 1.12-1.41; 139%-250% vs ≥251%: ARR, 1.19; 95% CI, 1.03-1.37); HIV-negative patients prescribed PrEP (ARR, 1.25; 95% CI, 1.00-1.55) vs HIV-negative patients not prescribed PrEP; and number of cohort years (ARR, 1.06; 95% CI, 1.02-1.11). The risk of moderate-to-severe depressive symptoms was lower for older age group patients vs the youngest age group (31-40 vs 18-24: ARR, 0.77; 95% CI, 0.67-0.87; 41-50 vs 18-24: ARR, 0.81; 95% CI, 0.68-0.96; >51 vs 18-24: ARR, 0.64; 95% CI, 0.52-0.79); patients identifying as Asian (ARR, 0.70; 95% CI, 0.55-0.89) and Black (ARR, 0.79; 95% CI, 0.66-0.95) vs White; and for Callen-Lorde vs Fenway patients (ARR, 0.23; 95% CI, 0.19-0.29).

E-Value Sensitivity Analysis

In the sensitivity analysis, we calculated an E-value of 1.63 for the covariate-adjusted association between GAHT and moderate-to-severe depressive symptoms. This implies that an unmeasured or uncontrolled confounder would need to be associated with a 1.63-fold increased risk for moderate-to-severe depressive symptoms for confounding to completely explain the observed association.

Discussion

This study found that GAHT prescribed using an informed consent protocol was associated with significantly lower rates of PHQ depression morbidity over time in a diverse multisite cohort of TGD adult primary care patients. The TGD patients prescribed GAHT had a 15% lower risk of past 2-week moderate-to-severe depressive symptoms across follow-up, compared with those not prescribed GAHT. Corroborating prior evidence,[10-13] GAHT was associated with lower rates of moderate-to-severe depressive symptoms for TGD patients. The mechanisms through which GAHT improves depressive symptoms are likely biopsychosocial for TGD people,[35] including physiologic changes in the hormone milieu, reductions in gender dysphoria, and increases in gender congruence impacting social functioning.[15] These findings highlight that the provision of GAHT in primary care settings using an informed consent model may improve moderate-to-severe depressive symptoms for TGD patients.

More than 1 in 10 (15.3%) of TGD patients met criteria for moderate-to-severe depressive symptoms using the PHQ, higher than US national probability sample data demonstrating 6.8% of adults experienced PHQ depression in 2019.[4] The differential distribution of PHQ depression found in the current study, with a higher prevalence in younger age groups and lower socioeconomic groups, and a lower prevalence and Black individuals, is consistent with national statistics for depression[1,3] and depressive symptoms in the general US population.[4] Yet, the consideration of gender identity differences in PHQ depression and data related to GAHT prescribing is a novel

Table 2. The LEGACY Cohort: Gender-Affirming Hormone Prescription and PHQ Depression in Last 2 Weeks by Year, 2016-2019[a]

		No. (%)			
		Exposure		Outcome	
Cohort years	Total No. in cohort	Prescribed gender-affirming hormones	Not prescribed gender-affirming hormones	PHQ depression	No PHQ depression
2016	3592	3034 (84.5)	558 (15.5)	551 (15.3)	3041 (84.7)
2017	3378	3117 (92.3)	261 (7.7)	493 (14.6)	2885 (85.4)
2018	3410	3191 (93.6)	219 (6.4)	444 (13.0)	2966 (87.0)
2019	3358	3164 (94.2)	194 (5.8)	387 (11.5)	2971 (88.5)

Abbreviation: PHQ, Patient-Health Questionnaire.

[a] Depression was defined as having a PHQ-2 score of 3 points or more (for Callen-Lorde patients) or a PHQ-9 score of 10 points or more (for Fenway Health patients) indicative of clinically significant symptoms of distress in the past 2 weeks (moderate-to-severe depressive symptoms).

Downloaded from jamanetwork.com by guest on 03/18/2025

contribution of this study, given that many public health surveillance efforts do not routinely collect TGD-specific information.[3,4] Indeed, transgender women and patients who were nonbinary assigned female sex at birth had a significantly increased risk of moderate-to-severe depressive symptoms compared with transgender men in the cohort.

Furthermore, moderate-to-severe depressive symptoms were investigated alongside HIV morbidity and prevention-effective adherence, contextualizing mental health within the HIV

Table 3. Models of Patients' Risk of PHQ for Depression, 2016-2019[a]

Variable	Bivariate models, RR (95% CI)	Multivariable model, ARR (95% CI)
Exposure		
Prescribed gender-affirming hormones		
Yes	0.72 (0.62-0.83)	0.85 (0.75-0.98)
No	1 [Reference]	1 [Reference]
Covariates and confounders		
Site		
Callen-Lorde	0.22 (0.20-0.25)	0.22 (0.20-0.25)
Fenway Health	1 [Reference]	1 [Reference]
Age group, y		
18-24	1 [Reference]	1 [Reference]
25-30	0.88 (0.79-0.99)	0.90 (0.81-1.00)
31-40	0.69 (0.60-0.80)	0.77 (0.67-0.87)
41-50	0.71 (0.59-0.85)	0.81 (0.68-0.96)
≥51	0.64 (0.51-0.79)	0.64 (0.52-0.79)
Gender identity		
Transgender woman	0.98 (0.88-1.10)	1.18 (1.05-1.32)
Transgender man	1 [Reference]	1 [Reference]
Nonbinary AFAB	1.71 (1.48-1.97)	1.17 (1.01-1.35)
Nonbinary AMAB	1.51 (1.22-1.88)	1.05 (0.85-1.30)
Racial identity		
American Indian or Alaskan Native	0.56 (0.31-1.02)	0.90 (0.52-1.54)
Asian	0.62 (0.49-0.79)	0.70 (0.55-0.89)
Black or African American	0.50 (0.42-0.61)	0.79 (0.66-0.95)
Multiracial	0.99 (0.80-1.21)	1.01 (0.86-1.19)
Native Hawaiian or Other Pacific Islander	0.37 (0.20-0.70)	0.72 (0.40-1.30)
White	1 [Reference]	1 [Reference]
Another racial identity	1.65 (1.11-2.45)	0.96 (0.63-1.45)
Ethnicity		
Latinx or Hispanic	0.60 (0.52-0.71)	0.90 (0.76-1.07)
Not Latinx or Hispanic	1 [Reference]	1 [Reference]
Insurance		
Private insurance	1 [Reference]	1 [Reference]
Public insurance	1.06 (0.95-1.17)	1.35 (1.22-1.49)
Uninsured	0.98 (0.75-1.29)	1.05 (0.82-1.34)
Federal poverty level, %		
0-138	1.17 (1.03-1.34)	1.26 (1.12-1.41)
139-250	1.10 (0.94-1.28)	1.19 (1.03-1.37)
251	1 [Reference]	1 [Reference]
HIV risk category		
HIV-negative, not prescribed PrEP	1 [Reference]	1 [Reference]
HIV-negative, prescribed PrEP	1.56 (1.20-2.02)	1.25 (1.00-1.55)
Living with HIV, virally suppressed[b]	0.72 (0.53-0.99)	1.12 (0.83-1.51)
Living with HIV, not virally suppressed	0.81 (0.57-1.16)	1.32 (0.91-1.91)
Total years in cohort	0.96 (0.92-1.00)	1.06 (1.02-1.11)

Abbreviations: AFAB, assigned female sex at birth; AMAB, assigned male sex at birth; ARR, adjusted risk ratio; PHQ, Patient-Health Questionnaire; PrEP, preexposure prophylaxis; RR, risk ratio.

[a] Longitudinal bivariate and multivariable models adjusted for continuous observation year with 25 320 observations. Depression was defined as having a PHQ-2 score of 3 points or more (for Callen-Lorde patients) or a PHQ-9 score of 10 points or more (for Fenway Health patients) indicative of clinically significant symptoms of distress in the past 2 weeks (moderate-to-severe depressive symptoms).

[b] Virally suppressed is defined as having an HIV viral load count <200 copies/mL.

Downloaded from jamanetwork.com by guest on 03/18/2025

epidemiologic burden.[14] Patients who were HIV-negative who were prescribed PrEP had higher rates of moderate-to-severe depressive symptoms than HIV-negative patients not prescribed PrEP. These findings highlight the need for integrated primary care and mental health services, including interventions that combine GAHT with HIV prevention and care.[36,37] In this study, only 1.7% of HIV-negative patients were receiving PrEP, and 22.0% of people living with HIV were not virally suppressed, well above the goals for HIV prevention and treatment in the 2022-2025 National HIV/AIDS Strategy for the US.[38]

Strengths and Limitations

Cohort strengths included diversity in age, gender identity, race and ethnicity, socioeconomic characteristics, and HIV serostatus; patients from primary care rather than gender clinics; a high number of nonbinary-identified patients; and measurement of other TGD health inequities (eg, HIV acquisition and transmission). The longitudinal design allowed for the estimation of exposure-outcome associations across 48 months of follow-up, the EHR dataset was substantially larger than many TGD clinical samples, and the sites were clinical settings. Study results are particularly relevant in the contemporary US context for TGD people. In recent years, an unprecedented number of anti-TGD bills and laws have been introduced,[39] many of which limit access to or ban gender-affirming care (eg, 2 bills were introduced in 2018 pertaining to gender-affirming care vs 185 bills in 2023).[40] Furthermore, recent US representative data demonstrate the widening of mental health inequities for TGD vs cisgender adults between 2018 and 2022, as indicated by increases in the percentage of TGD adults compared with cisgender adults who report 14 days or more of poor mental health in the past month.[41]

These findings should be interpreted alongside several limitations. This was a cohort from FQHCs that specialize in providing GAHT through an informed consent model. The association of GAHT with moderate-to-severe depressive symptoms may, therefore, be confounded by the gender-affirming context in which GAHT prescribing occurred (eg, trained clinicians, correct name and pronoun use). The TGD patients in this clinic-based cohort may be healthier than those not receiving care (ie, clinic patient bias). The cohort was relatively young and based in urban settings, which may impact the representativeness of the findings. Similarly, regarding generalizability, FQHCs serve safety net patient populations and have more comprehensive behavioral health services than non-FQHC practices[42]; thus, TGD patients in this cohort may be more disenfranchised socioeconomically or have better access to behavioral health care than TGD patients in other primary care settings. Selection bias may also be a relevant limitation, since participants self-selected into GAHT; however, the nonrandomized design was ethically appropriate for the research question.[43] The study did not assess the duration of GAHT or uptake of gender-affirming surgery; future research is needed to longitudinally explore these issues. The outcome variable was the PHQ screener for depression—not a diagnostic tool. Furthermore, the PHQ-9 was used at one clinical site while the PHQ-2 was used at the other site, which could result in unmeasured site differences due to the use of different metrics. However, the PHQ is a practical and recommended screening tool routinely implemented in primary care settings that can identify patients who may need further clinical evaluation and mental health referrals.[5-7]

Additionally, several potential unmeasured confounders were not assessed in this study, such as psychotropic medication use or other treatments for depression, substance abuse, housing insecurity, lack of family support, or experiences of stigma. We were also unable to account for the possibility that patients prescribed GAHT may have had more regular engagement with health professionals than those not prescribed GAHT, which is a potential confounder given prior research reporting the health-promoting role of collaborative patient-clinician interactions.[44] We calculated an E-value of 1.63, which means that an unmeasured or uncontrolled confounder would need to be associated with a 1.63-fold increase in the risk of moderate-to-severe depressive symptoms to completely explain the observed association in this study. In our multivariable model, no variables were associated with an adjusted relative risk greater than 1.63 (the highest risk in magnitude was for

Downloaded from jamanetwork.com by guest on 03/18/2025

public insurance, which was 1.35); thus, our results suggest that the observed association of GAHT and moderate-to-severe depressive symptoms in this study may be unlikely entirely explained by uncontrolled confounding. In addition, future analyses are needed to build on and strengthen evidence in this study, such as restricting the cohort to patients with depression at baseline and examining for evidence of downward trajectories in depression rates over time, investigating longitudinal depressive symptom trajectories to assess symptom changes from baseline to follow-up for patients after GAHT initiation vs corresponding changes in patients not initiating GAHT, and exploring the rapidity of effects in terms of timing, onset, and patterns of change.

Conclusions

In this cohort study, being prescribed GAHT was associated with lower rates of moderate-to-severe depressive symptoms in this study of diverse TGD primary care patients from FQHCs. Findings support the mental health–promoting role of GAHT and contribute to the evidence base that GAHT is a medically necessary treatment that improves mental health outcomes longitudinally for TGD people for those who want it. Study findings highlight the importance of gender-affirming care models delivered in primary care settings, low-barrier access to GAHT for TGD patients, and integration of GAHT into primary care with other services to address TGD mental health inequities. Furthermore, the prevalence of moderate-to-severe depressive symptoms in this TGD cohort underscores the importance of universal screening for depression in this population accompanied by appropriate systems to ensure diagnosis, treatment, and follow-up for those with depressive symptoms. Tailored interventions for subgroups, such as younger TGD patients and transgender women, may be necessary to address subpopulations at the highest risk and enhance early detection, diagnosis, and referrals to treatment for individuals with moderate-to-severe depressive symptoms.

ARTICLE INFORMATION

Accepted for Publication: January 14, 2025.

Published: March 17, 2025. doi:10.1001/jamanetworkopen.2025.0955

Open Access: This is an open access article distributed under the terms of the CC-BY-NC-ND License, which does not permit alteration or commercial use, including those for text and data mining, AI training, and similar technologies. © 2025 Reisner SL et al. *JAMA Network Open*.

Corresponding Author: Sari L. Reisner, ScD, MA, Department of Epidemiology, University of Michigan School of Public Health; 1415 Washington Heights, SPH I 2649A, Ann Arbor, MI 48109 (sreisner@umich.edu).

Author Affiliations: Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor (Reisner, Pletta); Center for Social Epidemiology and Population Health, University of Michigan, Ann Arbor (Reisner); Eisenberg Family Depression Center, University of Michigan, Ann Arbor (Reisner); Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, Massachusetts (Reisner); The Fenway Institute, Fenway Health, Boston, Massachusetts (Reisner, Keuroghlian, Mayer, Potter, Hughto); Department of Psychiatry, Massachusetts General Hospital, Boston/Harvard Medical School, Boston (Keuroghlian); Department of Medicine, Beth Israel Deaconess Medical Center/Harvard Medical School, Boston, Massachusetts (Mayer, Potter); Family Community Medicine, University of California, San Francisco (Deutsch); Department of Behavioral and Social Sciences, Brown University School of Public Health, Providence, Rhode Island (Hughto); Department of Epidemiology, Brown University School of Public Health, Providence, Rhode Island (Hughto); Center for Health Promotion and Health Equity, Brown University, Providence, Rhode Island (Hughto); Callen-Lorde Community Health Center, New York, New York (Harris, Radix); Department of Epidemiology, Columbia University Mailman School of Public Health, New York, New York (Radix).

Author Contributions: Drs Reisner and Pletta had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Reisner, Pletta, Potter, Radix.

Acquisition, analysis, or interpretation of data: Reisner, Pletta, Keuroghlian, Mayer, Deutsch, Hughto, Harris, Radix.

Downloaded from jamanetwork.com by guest on 03/18/2025

Drafting of the manuscript: Reisner, Pletta, Harris.

Critical review of the manuscript for important intellectual content: All authors.

Statistical analysis: Reisner, Pletta.

Obtained funding: Reisner.

Administrative, technical, or material support: Reisner, Keuroghlian, Mayer, Deutsch, Potter, Hughto, Harris, Radix.

Supervision: Reisner, Mayer.

Conflict of Interest Disclosures: Dr Reisner reported receiving royalties from McGraw Hill for editing a textbook outside the submitted work. Dr Keuroghlian reported receiving a patent for editing with royalties paid from McGraw Hill for a textbook on transgender and gender diverse health care and a patent for editing with royalties paid from the American Psychiatric Association Publishing for a textbook on gender-affirming psychiatric care. No other disclosures were reported.

Funding/Support: The research reported in this publication was supported by the Patient-Centered Research Outcomes Institute (PCORI) under award AD-2017C1-6569 (Dr Reisner).

Role of the Funder/Sponsor: PCORI had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the PCORI.

Data Sharing Statement: See the Supplement.

Additional Contributions: We thank the following individuals and entities from Fenway Health for their contributions to LEGACY: Andrew Asquith, BA, for study coordination; Juwan Campbell, MS, for data management; Alex Gonzalez, MD, MPH, for clinical consultation on study protocol; Dana Pardee, MBA, for project management and oversight; Meg Quint, BS, for support with the literature review; the community advisory board and the scientific advisory board for feedback on the study protocol and implementation; and TGD patients at Fenway Health and Callen-Lorde for participating.

REFERENCES

1. Lee B, Wang Y, Carlson SA, et al. National, state-level, and county-level prevalence estimates of adults aged ≥18 years self-reporting a lifetime diagnosis of depression—United States, 2020. *MMWR Morb Mortal Wkly Rep*. 2023;72(24):644-650. doi:10.15585/mmwr.mm7224a1

2. Zhang Z, Jackson SL, Gillespie C, Merritt R, Yang Q. Depressive symptoms and mortality among US adults. *JAMA Netw Open*. 2023;6(10):e2337011. doi:10.1001/jamanetworkopen.2023.37011

3. Substance Abuse and Mental Health Services Administration (SAMHSA). Key substance use and mental health indicators in the United States: results from the 2021 National Survey on Drug Use and Health (HHS publication No. PEP22-07-01-005, NSDUH Series H-57). 2022. Accessed July 20, 2024. https://www.samhsa.gov/data/report/2021-nsduh-annual-national-report

4. Centers for Disease Control and Prevention (CDC). Symptoms of depression among adults: United States, 2019 (National Center for Health Statistics Data Brief No. 379). 2020. Accessed July 20, 2024. https://www.cdc.gov/nchs/products/databriefs/db379.htm

5. Siu AL, Bibbins-Domingo K, Grossman DC, et al; US Preventive Services Task Force (USPSTF). Screening for depression in adults: US Preventive Services Task Force recommendation statement. *JAMA*. 2016;315(4):380-387. doi:10.1001/jama.2015.18392

6. Barry MJ, Nicholson WK, Silverstein M, et al; US Preventive Services Task Force. Screening for depression and suicide risk in adults: US Preventive Services Task Force recommendation statement. *JAMA*. 2023;329(23):2057-2067. doi:10.1001/jama.2023.9297

7. Reynolds CF III, Frank E. US Preventive Services Task Force recommendation statement on screening for depression in adults: not good enough. *JAMA Psychiatry*. 2016;73(3):189-190. doi:10.1001/jamapsychiatry.2015.3281

8. Reisner SL, Poteat T, Keatley J, et al. Global health burden and needs of transgender populations: a review. *Lancet*. 2016;388(10042):412-436. doi:10.1016/S0140-6736(16)00684-X

9. Scheim AI, Rich AJ, Zubizarreta D, et al. Health status of transgender people globally: a systematic review of research on disease burden and correlates. *PLoS One*. 2024;19(3):e0299373. doi:10.1371/journal.pone.0299373

10. Baker KE, Wilson LM, Sharma R, Dukhanin V, McArthur K, Robinson KA. Hormone therapy, mental health, and quality of life among transgender people: a systematic review. *J Endocr Soc*. 2021;5(4):bvab011. doi:10.1210/jendso/bvab011

Downloaded from jamanetwork.com by guest on 03/18/2025

11. White Hughto JM, Reisner SL. A systematic review of the effects of hormone therapy on psychological functioning and quality of life in transgender individuals. *Transgend Health*. 2016;1(1):21-31. doi:10.1089/trgh.2015.0008

12. Doyle DM, Lewis TOG, Barreto M. A systematic review of psychosocial functioning changes after gender-affirming hormone therapy among transgender people. *Nat Hum Behav*. 2023;7(8):1320-1331. doi:10.1038/s41562-023-01605-w

13. van Leerdam TR, Zajac JD, Cheung AS. The effect of gender-affirming hormones on gender dysphoria, quality of life, and psychological functioning in transgender individuals: a systematic review. *Transgend Health*. 2023;8(1):6-21. doi:10.1089/trgh.2020.0094

14. Stutterheim SE, van Dijk M, Wang H, Jonas KJ. The worldwide burden of HIV in transgender individuals: an updated systematic review and meta-analysis. *PLoS One*. 2021;16(12):e0260063. doi:10.1371/journal.pone.0260063

15. Coleman E, Radix AE, Bouman WP, et al. Standards of care for the health of transgender and gender diverse people, version 8. *Int J Transgend Health*. 2022;23(suppl 1):S1-S259. doi:10.1080/26895269.2022.2100644

16. Reisner SL, Bradford J, Hopwood R, et al. Comprehensive transgender healthcare: the gender affirming clinical and public health model of Fenway Health. *J Urban Health*. 2015;92(3):584-592. doi:10.1007/s11524-015-9947-2

17. Asquith A, Sava L, Harris AB, Radix AE, Pardee DJ, Reisner SL. Patient-centered practices for engaging transgender and gender diverse patients in clinical research studies. *BMC Med Res Methodol*. 2021;21(1):202. doi:10.1186/s12874-021-01328-4

18. LeBlanc M, Radix A, Sava L, et al. "Focus more on what's right instead of what's wrong:" research priorities identified by a sample of transgender and gender diverse community health center patients. *BMC Public Health*. 2022;22(1):1741. doi:10.1186/s12889-022-14139-z

19. Reisner SL, Deutsch MB, Mayer KH, et al. Longitudinal cohort study of gender affirmation and HIV-related health in transgender and gender diverse adults: the LEGACY Project protocol. *JMIR Res Protoc*. 2021;10(3):e24198. doi:10.2196/24198

20. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61(4):344-349. doi:10.1016/j.jclinepi.2007.11.008

21. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-613. doi:10.1046/j.1525-1497.2001.016009606.x

22. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care*. 2003;41(11):1284-1292. doi:10.1097/01.MLR.0000093487.78664.3C

23. US Department of Health and Human Services (HHS). Federal poverty level (FPL). 2022. Accessed July 10, 2024. https://www.healthcare.gov/glossary/federal-poverty-level-fpl/

24. Haberer JE, Bangsberg DR, Baeten JM, et al. Defining success with HIV pre-exposure prophylaxis: a prevention-effective adherence paradigm. *AIDS*. 2015;29(11):1277-1285. doi:10.1097/QAD.0000000000000647

25. Panel on Antiretroviral Guidelines for Adults and Adolescents. Guidelines for the use of antiretroviral agents in adults and adolescents with HIV. 2024. Accessed July 10, 2024. https://clinicalinfo.hiv.gov/en/guidelines/adult-and-adolescent-arv

26. Haneuse S, Arterburn D, Daniels MJ. Assessing missing data assumptions in EHR-based studies: a complex and underappreciated task. *JAMA Netw Open*. 2021;4(2):e210184. doi:10.1001/jamanetworkopen.2021.0184

27. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res*. 2007;16(3):219-242. doi:10.1177/0962280206074463

28. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45(3):1-67. doi:10.18637/jss.v045.i03

29. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*. 1986;42(1):121-130. doi:10.2307/2531248

30. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics*. 1988;44(4):1049-1060. doi:10.2307/2531734

31. Rubin D. *Multiple Imputation for Nonresponse in Surveys*. Wiley-Interscience; 2004.

32. Neuhaus J, Kalbfleisch J, Hauck W. A comparison of cluster-specific and population-averaged approaches for analyzing correlated binary data. *Int Stat Rev*. 1991;59(1):25-35. doi:10.2307/1403572

Downloaded from jamanetwork.com by guest on 03/18/2025

33. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-Value. *Ann Intern Med*. 2017;167(4):268-274. doi:10.7326/M16-2607

34. Mathur MB, Ding P, Riddell CA, VanderWeele TJ. Web site and R package for computing E-values. *Epidemiology*. 2018;29(5):e45-e47. doi:10.1097/EDE.0000000000000864

35. Engel GL. The clinical application of the biopsychosocial model. *Am J Psychiatry*. 1980;137(5):535-544. doi:10.1176/ajp.137.5.535

36. Reisner SL, Radix A, Deutsch MB. Integrated and gender-affirming transgender clinical care and research. *J Acquir Immune Defic Syndr*. 2016;72(Suppl 3)(suppl 3):S235-S242. doi:10.1097/QAI.0000000000001088

37. Poteat T, Scheim A, Xavier J, Reisner S, Baral S. Global epidemiology of HIV infection and related syndemics affecting transgender people. *J Acquir Immune Defic Syndr*. 2016;72(Suppl 3)(suppl 3):S210-S219. doi:10.1097/QAI.0000000000001087

38. The White House. National HIV/AIDS strategy for the United States 2022–2025. 2021. Accessed July 10, 2024. https://files.hiv.gov/s3fs-public/NHAS-2022-2025.pdf

39. Movement Advancement Project. Equality maps: Snapshot: LGBTQ equality by state. 2024. Accessed September 8, 2024. https://www.mapresearch.org/equality-maps/

40. Trans Legislation Tracker. Tracking the rise of anti-trans bills in the US. 2024. Accessed September 9, 2024. https://translegislation.com/

41. Liu M, Patel VR, Reisner SL, Keuroghlian AS. Health status and mental health of transgender and gender-diverse adults. *JAMA Intern Med*. 2024;184(8):984-986. doi:10.1001/jamainternmed.2024.2544

42. Lewis VA, Spivack S, Murray GF, Rodriguez HP. FQHC designation and safety net patient revenue associated with primary care practice capabilities for access and quality. *J Gen Intern Med*. 2021;36(10):2922-2928. doi:10.1007/s11606-021-06746-0

43. Deutsch MB, Radix A, Reisner S. What's in a guideline? developing collaborative and sound research designs that substantiate best practice recommendations for transgender health Care. *AMA J Ethics*. 2016;18(11):1098-1106. doi:10.1001/journalofethics.2016.18.11.stas1-1611

44. Rathert C, Mittler JN, Vogus TJ, Lee YSH. Better outcomes through patient-provider therapeutic connections? an exploratory study of proposed mediating variables. *Soc Sci Med*. 2023;338:116290. doi:10.1016/j.socscimed.2023.116290

SUPPLEMENT.

Data Sharing Statement

Downloaded from jamanetwork.com by guest on 03/18/2025